DOCUMENT RESUME

ED 333 016 TM 016 465

AUTHOR Green, Kathy E.; Kluever, Raymond C.

TITLE Component Identification and Item Difficulty of Raven's Matrices Items.

PUB DATE Apr 91

NOTE 17p.; Paper presented at the Annual Meeting of the National Council on Measurement in Education (Chicago, IL, April 4-6, 1991).

PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.

DESCRIPTORS Academically Gifted; *Children; Comparative Testing; *Difficulty Level; Elementary Education; *Elementary School Students; Item Response Theory; Mathematical Models; Predictive Measurement; Preschool Education; *Psychological Testing; Regression (Statistics); Test Construction; *Test Items

IDENTIFIERS Rasch Model; *Ravens Coloured Progressive Matrices; *Standard Progressive Matrices

ABSTRACT

Item components that might contribute to the difficulty of items on the Raven Colored Progressive Matrices (CPM) and the Standard Progressive Matrices (SPM) were studied. Subjects providing responses to CPM items were 269 children aged 2 years 9 months to 11 years 8 months, most of whom were referred for testing as potentially gifted. A second sample containing 147 seventh-grade students, drawn from J. K. Gallini's study in 1983, was used to assess the utility of the equation developed using the first item sample. CPM item characteristics were defined and rated. Rasch item difficulties were used as the dependent variable with misfitting items omitted. All 15 item characteristics were entered in a regression equation using forced entry (multiple "R" of 0.90) and stepwise entry (multiple "R" of 0.88). When the same predictors were used with SPM items, the multiple "R" was 0.69. The poorest prediction occurred for items containing characteristics (such as line thickness) that were not captured by the coding system. The best prediction occurred for items in which the orientation of the figure or options was a crucial feature. Results are discussed with regard to psychological processes and use of item characteristics to create new test items. Two sample test items are included, and two tables and an appendix present data on item difficulties. (SLD)

COMPONENT IDENTIFICATION AND ITEM DIFFICULTY OF RAVEN'S MATRICES ITEMS

Kathy E. Green and Raymond C. Kluever

University of Denver

ABSTRACT

The purpose of this study was to identify and test item characteristics that predict the difficulty of Raven's Colored and Standard Progressive Matrices items. Colored Progressive Matrices item characteristics were defined and rated; Rasch item difficulties were used as the dependent variable with misfitting items omitted. The multiple R was .90 (.88 using stepwise prediction). When the same predictors were used with Standard Progressive Matrices items, the multiple R was .69. Results are discussed with respect to psychological processes and to using item characteristics to create new test items.

Paper presented at the Annual Meeting of the National Council on Measurement in Education, Chicago, April 4-6, 1991.

INTRODUCTION

Figural reasoning items such as those developed by Raven (1965, 1985) have proven useful as nonverbal measures of fluid or analytic intelligence ("g"). While matrices tests have been used for years and while a number of researchers have examined the factor structure of these tests, little attention has been paid to empirical examination of the item variations contributing to item difficulty. The purpose of this study was to identify item components that might contribute to item difficulty and then to assess which components are predictive of empirical item difficulty. The prediction equation constructed was then used to predict difficulty of a second set of matrices items.

The Raven Colored Progressive Matrices (CPM) and Standard Progressive Matrices (SPM) have been used as measures of a unitary trait, although factor analyses suggest that 2-4 factors are necessary to explain item intercorrelations (e.g., Carlson & Jensen, 1980; Schmidtke & Schaller, 1980). Analyses indicated that while a multifactor solution was needed, item responses yielded adequate fit to a logistic model (the Rasch model: Green & Kluever, 1991). High internal consistency coefficients have also been reported (e.g., Court & Raven, 1982; Raven, Court & Raven, 1986). If all test items are reflective of a single latent trait, it should be possible to identify ways in which items differ that make some items systematically easier and others harder to answer. This process is termed component identification. Identification of components would prove useful in construction of additional test items and in furtherance of our understanding of the construct being measured.

Suggestions of both method of analysis and of potential components may be obtained from previous work with Knox Cube Block Test items, paper-folding items, and progressive matrices items (Carpenter, Just, & Shell, 1990; Gallini, 1983; Green & Smith, 1987; Smith & Green, 1985; Ward & Fitzpatrick, 1973). Some of these item types differ from matrices items but all are non-verbal. Analyses of verbal analogies items have also been performed but these results seem less relevant to the analysis of nonverbal items.

The number and complexity of components determine the difficulty of an item. Potential components relevant to matrices items defined by previous work include variables such as symmetry vs. assymmetry, vertical or horizontal vs. diagonal axes, straight lines vs. curved lines, size of cell attributes, number of dimensions of variation (e.g., different line widths, different shapes), proportionate vs. disproportionate change of size, shading, number of colors, intersection vs. union of dimensions, rotation of elements, and reflection of elements. Number, orientation, and figure type are the problem descriptors used by Carpenter et al. (1990) in their analysis of cognitive processes used by high and average performers on matrices items. More difficult problems involve more figural elements and/or more complex combination rules (Ward & Fitzpatrick, 1973). If multiple rules are needed in problem solution, cognitive management of rule construction and execution is taxed as well as the mental processes used to construct the rules. Carpenter et al. (1990) suggest that individual performance differences reflect abilities to generate and maintain problem solving goals in working memory. Mulholland, Pellegrino, and Glaser (1980) found that errors and response times increased when the number of operations needed to solve geometric analogies increased. They also attributed this performance decrement to an increased burden on working memory created by the need to track more elements and transformations.

The skill with which individuals process information is dependent on variables such as the kind(s) of cognitive processes involved, the nature of the content, the complexity of processing required, and one's previous experience with the task. Basic processing models involving variations of the input-process-output systems are common in the literature. One's experience with the content of the material is reflected in some processing models. Other models are more descriptive of the nature of the material and some models are more concerned with the cognitive complexity required to solve a problem.

The Structure of Intellect (SOI) model (Guilford, 1959) lends itself as a framework to systematic analysis of the content and processes involved in solving matrices problems. Other relevant models are simultaneous-sequential processing models and Horn and Cattell's (1966) model of cognitive processing. Item difficulty and item characteristics are related in this paper to these models as indications of the cognitive processes involved, the nature of the content, and its complexity.

METHOD

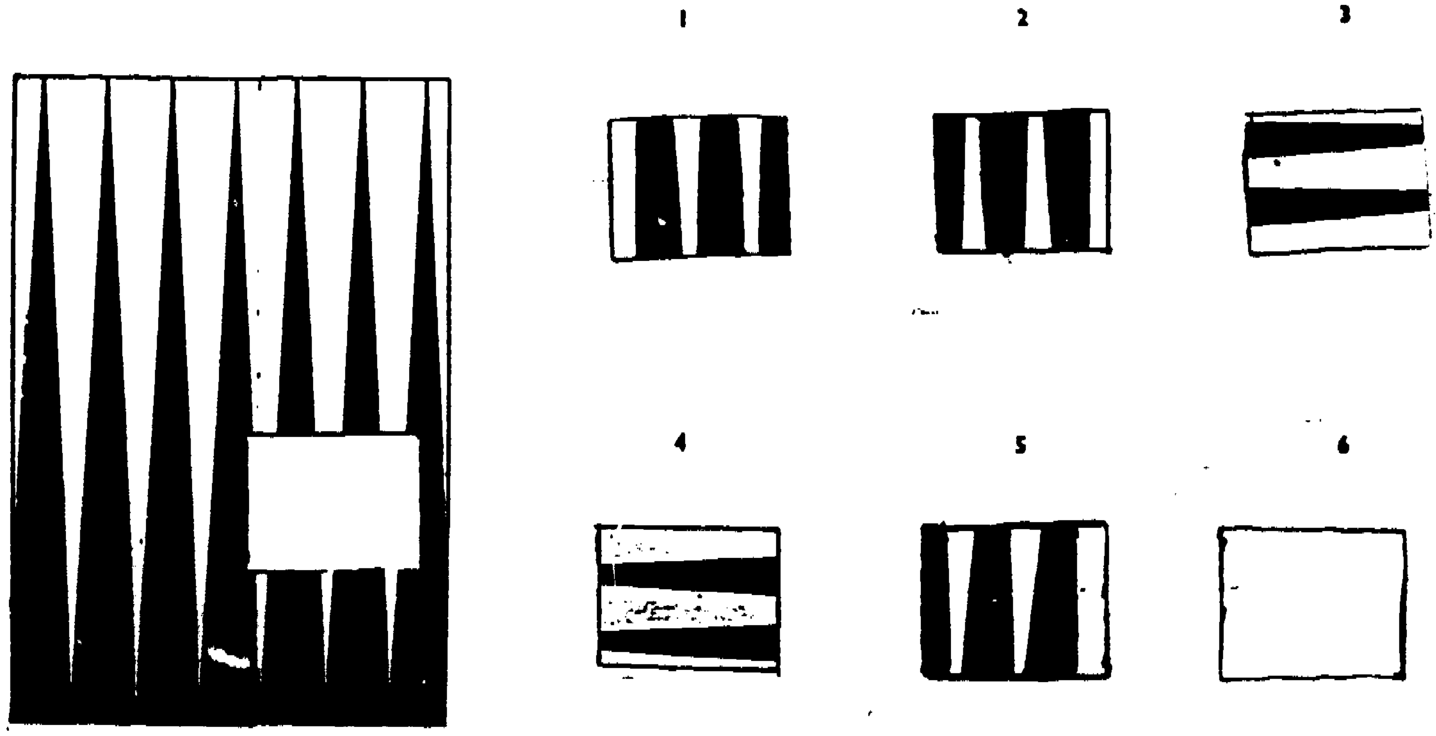
Subjects providing responses to CPM items were 269 2-9 to 11-8 year-old children seen at a University Assessment Center for individual testing through 1989. The majority of these children were referred for testing as potentially gifted. Responses were fit to a linear logistic model using BICAL (Wright, Mead, & Bell, 1980). Two of the 36 CPM items misfit (both between and total fit >3.00) and so were dropped from subsequent analyses. Logit item difficulty obtained from BICAL was used as the dependent variable; the effective _n_ for this study was 34. A second sample was obtained from data reported by Gallini (1983) who gave 30 of the 60 Standard Progressive Matrices items to 147 seventh-grade students from an urban middle school in the Southeast. Four of these 30 items were answered correctly by all students so logit difficulties of one less than the lowest logit value were assigned to these items. Three other items misfit and were dropped from the analysis leaving an effective _n_ of 27 for this sample. The second sample was used to assess the utility of the equation developed using the first item sample.

Analyses were conducted using estimated Rasch item difficulties regressed on component frequencies. Item calibration using the Rasch model provides a means to evaluate unidimensionality. Regression analyses are of limited value unless the dependent variable assesses a single trait (in this case, item difficulty). Modeling performance on a set of items that are not well-defined can be pointless. Thus, Rasch analysis with removal of misfitting items was used to refine the dependent measure, to provide reasonable assurance of unidimensionality. Relationships among predictor variables was also assessed since the presence of multicollinearity leads to varied interpretation of results.

CPM items consist of an item "stem" which contains a figural display with a missing piece, and 6 or 8 response options, one of which completes the figural pattern. Stem characteristics assessed were: vertical/horizontal orientation vs. other, symmetrical vs. asymmetrical, progression vs. not, number of dimensions in the pattern, straight lines vs. curved lines, number of lines or solids, density of design, and color vs. black and white. Number of dimensions and elements were coded a 0-3; all others were coded as 0-1. Response option characteristics assessed were: number of distinct options (2-8), options contain progression, rotation, reflection (0-1 for each), number of directions of options (e.g., horizontal, vertical, diagonal; 1-3), number of elements in the design (1-3), and reversal between foreground and background (0-1). Characteristics were rated independently by the two authors; disagreements occurred for less clearly specified characteristics such as "density of design." Disagreements were either resolved or the category redefined. Figure 1 illustrates item component categorization for two hypothetical items.

EXAMPLES OF ITEM CHARACTERISTIC CODING

<u>Item 1</u>



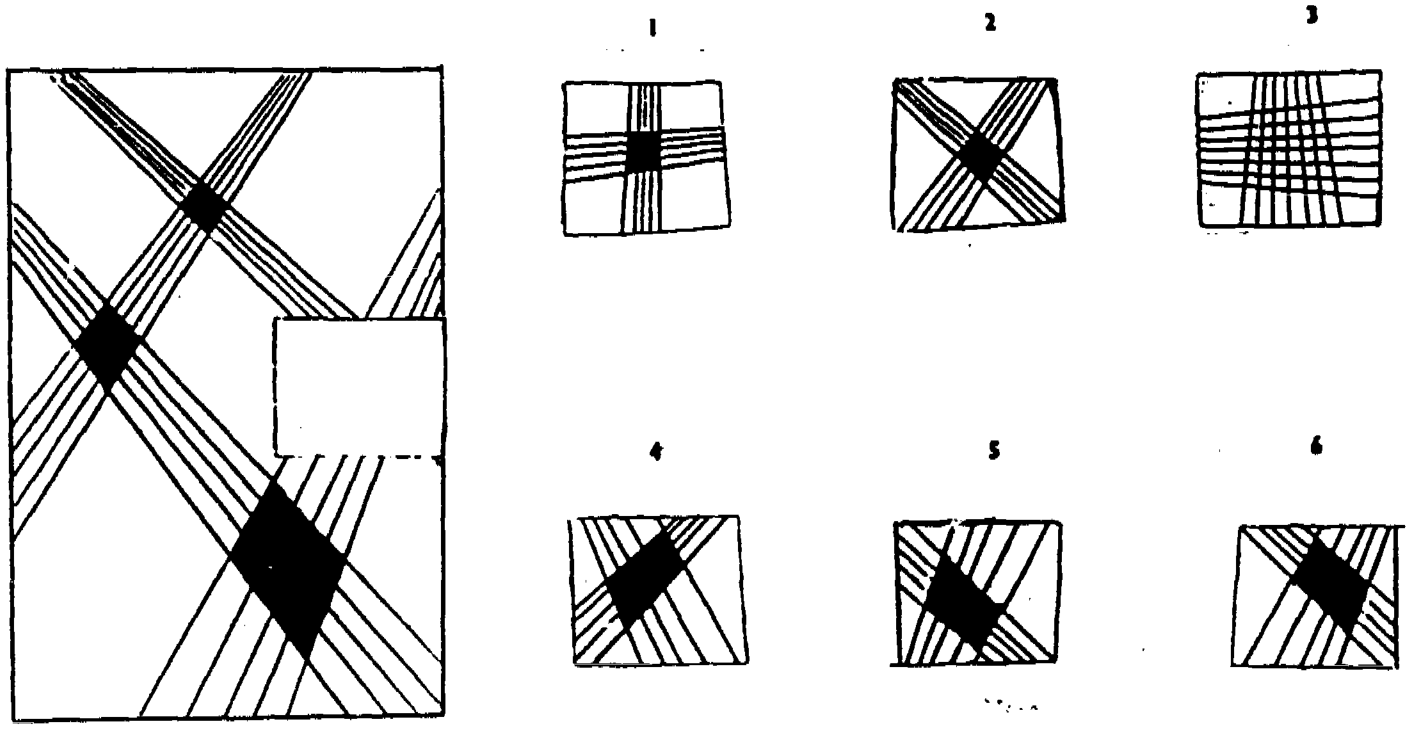
Stem Variables:

Vertical/horizontal orientation: 0-yes
Symmetrical: 0-yes
Progression: 0-no progression
dimensions in pattern: 1-pattern only
Straight lines vs. curved: 0-straight
elements: 1-lines
Color vs. black/white: 1-black/white

Option Variables:

distinct options: 5
Options contain progression: 0-no
rotation: 1-yes
reflection: 1-yes (e.g., 1 and 5)
option directions: 3 (vertical, horizontal, diagonal)
Reversal between foreground and background: 0-no

Item 2



Stem Variables:

Vertical/horizontal orientation: 1-diagonal
Symmetrical: 1-no
Progression: 1-yes (angle of line separation increases)
dimensions in pattern: 2-size, orientation
Straight lines vs. curved: 0-straight
elements: 2 (lines, solids)
Color vs. black/white: 1-black/white

Option Variables:

distinct options: 6
Options contain progression: 1-yes (e.g., 2 and 6)
rotation: 1-yes (e.g., 4 and 5)
reflection: 1-yes (e.g., 4 and 6)
option directions: 2 (vertical, diagonal)
Reversal between foreground and background: 0-no

RESULTS

Table 1 presents the zero-order correlations of all characteristics with item difficulty for CPM items. Item characteristics were multicollinear. Table 2 presents significant ($p<.05$) inter-characteristic correlations for CPM items. CPM items were used as the basis for construction of a regression function. All 15 item characteristics were entered in a regression equation using two methods: forced entry and stepwise entry. With forced entry, the multiple R was .90. With stepwise regression, number of distinct options, reflection of one or more options, number of dimensions in the stem, and number of directions of options contributed significantly ($p<.05$) to prediction for a multiple R of .88. The R^2 adjusted for number of cases was .74. Actual and predicted item difficulties for CPM items are provided in Appendix A as is the standardized regression function.

SPM item difficulty was predicted using the four characteristics identified as significant predictors of CPM item difficulty. The multiple R was .69. Actual and predicted item difficulties are listed in Appendix A.

To assess the effect of nonlinearity, the squared and cubed values of components were added to prediction equations. This resulted in a small increase in prediction for CPM items (.88 to .91) and in a similar increase in prediction for SPM items (.69 to .71). When higher order terms were added, the variable representing number of options dropped out of the function.

Table 1

Correlation of Item Characteristics with Item Difficulty

Characteristic	Correlation-CPM	SPM
Orientation	01	
Symmetry	61	
Progression	17	
Number of Dimensions	66	61
Lines	22	
Lines-Solids	68	
Color	46	
Progression	-18	
Rotation	28	
Reflection	34	31
Number of Options	48	52
Density	-02	
Number of Directions of Options	19	19
Reversal of Foreground and Background	02	
Number of Elements	22	

Table 2

Item Characteristic Intercorrelations-CPM

	C6	C7	C8	C13	C14	C16
C2: Symmetry	78	76	53			
C4: Number of Dimensions		82	69			
C6: Lines-Solids		61		38		41
C7: Color				45		
C9: Rotation					34	
C13: Density of Design					-48	

C8: Progression; C14: Number of Directions of Options; C16: Number of Elements in Design

Note: Only item characteristics that were significantly correlated ($p<.05$) are listed.

The two items from each test for which prediction was the poorest and best were reviewed. The poorest prediction occurred for items containing characteristics (such as line thickness) that were not captured by the coding system. The best prediction occurred for items in which orientation of figure or of options was a crucial feature.

DISCUSSION

Only tentative conclusions may be drawn from the results of this study. Some variables that were found to be significant predictors were correlated with other item characteristics. This collinearity makes determination of which unique characteristics predict item difficulty problematic.

Orientation of options (rotation and reflection) were both significantly related to item difficulty with reflection somewhat more highly correlated. Many progressive matrices items involve orientation as well as design matching and number. Items tend to be more difficult when several options are identical in shape/number but are mirror images of each other. Rotation of an option away from a vertical/horizontal orientation does not seem to pose as difficult a problem as reversal in the same plane. For Raven's items, reflection seems to require a finer discrimination than rotation. Rotation for these items is likely to involve a disturbance to the orientation of the figure which is relatively easy to identify as incorrect.

Number of directions of options also was a significant predictor for CPM items. This variable also assessed spatial orientation. This variable did not significantly add to prediction for SPM items.

Number of distinct options was a significant predictor for CPM items. All CPM items have six options and only two of the 30 items have less than six unique options. These are the second and third items on the test and are also among the easiest test items. SPM items have six or eight options; again, only two items have options that are not unique and these items are the easiest. This variable served to identify extremely easy items.

Number of dimensions in the stem assessed the number of different figural elements that needed to be considered in problem solution. Possible elements were figural match, orientation, size, and number. Items varying on more than one dimension were more difficult. This variable was a measure of item complexity and was most predictive of difficulty for both CPM and SPM items. This variable is similar to number of transformations which previous researchers have also found to be predictive of item difficulty.

While analysis of components does not describe the elementary mental processes necessary to problem solution, we propose congruity with certain processing models. Processing models may include components such as the ability to attend, to encode information, short- and long-term retention, the ability to retrieve infromation from storage, certain cognitive processing skills, and an output component. Task, content, and situational characteristics are important factors to consider in analyzing a specific processing requirement.

Among models of cognitive processing, the Horn and Cattell (1966) theory includes fluid and crystallized ability representing the processing of new versus more familiar material. Success with components of the CPM items probably places a high premium on fluid abilities since solving matrices tasks is not common. At best, the analysis leading to the correct choice for stimulus completion may parallel certain common everyday discrimination tasks but the CPM content is certainly unique.

Analysis of this fluid ability for solving the CPM problems can be in the context of a simultaneous and sequential processing scheme proposed by several psychologists (Das, Kirby, & Jarman, 1975; Luria, 1973). Solution of the CPM problems seems to require good simultaneous processing ability of a gestalt-like configuration with differing components contributing to processing load as identified in this study.

A model displaying greater detail of processing skills is Guilford's Structure of Intellect (SOI) Model (1967). The Raven designs probably represent the Figural Content as defined in the model. Among the five Operations described in the model, Cognition and Evaluation abilities appear to be most representative of the kind of processing required to solve the CPM problems. The Products of this model ranging from unitary components to complex abstracting abilities requiring the reconfiguration of the stimulus material are also evident in the Raven problem-solving requirements.

Both the SOI model and the simultaneous processing-sequential processing model have published tests based on those concepts. The value of the component analysis in this study is the opportunity it presents for construction of items that utilize component characteristics in revisions of these tests to tap a defined range of processing abilities. This has implications for test construction and for test interpretation where manuals could provide guidelines for this kind of interpretation.

There are several limitations of this study. First, all features of items were not assessed nor were problem-solving process variables assessed. Only fairly obvious observable features were included in the analysis. Second, the regression analysis assumes that a linear combination of variables predicts item difficulty. Even if the appropriate components have been identified, they may relate to other components and to item difficulty in a nonlinear fashion. Third, individual children may use different strategies in solving matrices items. The analyses performed implicitly assume that all subjects use similar strategies and attend to similar aspects of problems. This assumption may be overly simplistic. Finally, the size of the samples used in calibrating items were smaller than those desirable to establish highly reliable item statistics.

REFERENCES

Carpenter, P.A., Just, M.A., & Shell, P. (1990). What one intelligence test measures: A theoretical account of the processing in the Raven Progressive Matrices Test. _Psychological Review_, _97_(3), 404-431.

Carlsen, J.S., & Jensen, C.M. (1980). The factorial structure of the Raven Coloured Progressive Matrices Test: A reanalysis. _Educational and Psychological Measurement_, _40_, 1111-1116.

Das, J.P., Kirby, J.R., & Jarman, R.F. (1975). Simultaneous and successive synthesis: An alternative model for cognitive abilities. _Psychological Bulletin_, _82_, 87-103.

Gallini, J.K. (1983). A Rasch analysis of Raven item data. _Journal of Experimental Education_, _52_, 27-32.

Green, K.E., & Kluever, R.C. (1991). Structural properties of Raven's Colored Progressive Matrices for a sample of gifted children. _Perceptual and Motor Skills_, _72_, 59-64.

Green, K.E., & Smith, R.M. (1987). A comparison of two methods of decomposing item difficulties. _Journal of Educational Statistics_, _12_, 369-381.

Guilford, J.P. (1959). The three faces of intellect. _The American Psychologist_, _14_, 469-479.

Horn, J.L., & Cattell, R.B. (1966). Refinement and test of the theory of fluid and crystallized intelligence. _Journal of Educational Psychology_, _57_, 253-370.

Luria, A.R. (1973). _The working brain: An introduction to neuropsychology_. London: Penguin Books.

Mulholland, T.M., Pellegrino, J.W., & Glaser, R. (1980). Components of geometric analogy solution. _Cognitive Psychology_, _12_, 252-284.

Raven, J.C. (1965). _Guide to using the Coloured Progressive Matrices Sets A, Ab, B._ Dumfries, Scotland: Grieve.

Raven, J.C. (1985). _Standard Progressive Matrices Sets A,B,C,D, and E._ London: H.K. Lewis & Co.

Raven, J.C., Court, J.H., & Raven, J. (1977). _Coloured Progressive Matrices_. London: H.K. Lewis.

Schmidtke, A., & Schaller, S. (1980). Comparative study of factor structure of Raven's Coloured Progressive Matrices. _Perceptual and Motor Skills_, _51_, 1244-1246.

Smith, R.M., & Green, K.E. (1985). Components of difficulty in paper-folding tests. Paper presented at the annual meeting of the American Educational Research Association, Chicago.

Ward, J., & Fitzpatrick, T.F. (1973). Characteristics of matrices items. Perceptual and Motor Skills, 36, 987-993.

Wright, B.D., Mead, R.J., & Bell, S.R. (1980). BICAL: Calibrating items with the Rasch model (Research Memorandum 23C). Chicago: University of Chicago, Statistical Laboratory, Department of Education.

APPENDIX A

Actual and Predicted Item Difficulties for CPM and SPM Items

	Difficulty			Difficulty	
CPM Item	Logit	Predicted	SPM Item	Logit	Predicted
RA1	-2.83	-1.87	RA1	-5.16	-2.61
RA2	-4.66	-4.56	RA2	-5.16	-5.18
RA3	-3.97	-4.10	RA4	-3.00	-2.86
RA4	-3.40	-2.71	RA5	-.66	-2.86
RA5	-2.59	-2.71	RA6	1.13	-2.36
RA6	-2.11	-1.03	RA7	-5.16	-2.36
RA7	.23	-1.03	RA8	-5.16	-.94
RA8	-.05	.43	RB1	-2.43	-2.86
RA9	.30	.99	RB2	-.02	-2.61
RA10	.76	.15	RB3	-.74	1.04
RA11	2.51	.43	RB6	1.71	1.04
RA12	3.23	2.45	RC1	-2.21	-2.88
RAB1	-2.83	-1.03	RC2	-2.68	-2.88
RAB2	-1.35	-1.03	RC3	-.32	-2.88
RAB3	-1.32	.43	RC4	-.20	-2.88
RAB4	.21	.15	RC5	1.56	1.42
RAB5	.21	.99	RC6	1.37	.22
RAB6	1.04	-1.87	RD1	-3.44	-1.69
RAB7	.54	.99	RD2	-2.21	-1.69
RAB8	1.88	2.11	RD3	-.59	-1.69
RAB9	1.74	.15	RD5	.94	-2.88
RAB11	1.16	.15	RD6	2.90	-2.88
RB1	-4.66	-2.71	RE1	.03	1.42
RB2	-.81	-1.87	RE2	.56	.22
RB3	-.30	-.69	RE3	1.97	1.67
RB4	-.19	.15	RE4	2.02	2.13
RB5	1.02	.43	RE5	4.05	1.88
RB6	.99	1.61			
RB7	1.23	1.61			
RB8	2.73	3.91			
RB9	2.37	2.73			
RB10	1.83	1.89			
RB11	2.80	2.73			
RB12	4.26	2.73			

CPM: Z_y = .17(#options) + .43(reflection) + .75(#dimensions) + .29(# option directions) + e

SPM: Z_y = .31(#options) + .18(reflection) + .39(#dimensions) + .07(# option directions) + e